

# Gunabalan G

<https://github.com/gunabalan-0411> | <https://leetcode.com/u/gunabalan0411/> | [gunabalan0411@gmail.com](mailto:gunabalan0411@gmail.com)

## Experience – 2.6 Years

### HP Inc

*Data Scientist 2*

Bangalore, India

Apr 2023 – Present

- **Developed a scalable Full Stack AI** Chat Application, "**AI Analyst**," from architecture to deployment, reducing API usage costs from **\$13,000 to \$235** annually.
- **Built memory-optimized** Power BI dashboards.
- **Enhanced LLM models** with faster training and highly **scalable inference** using **parallel GPUs** for specific business use cases, creating a multi-level layered insight extractor.

### HP Inc

*Data Scientist*

Bangalore, India

Sep 2021 – Apr 2023

- **Automated end-to-end business intelligence** reports, from statistical analysis to PowerPoint presentations and email notifications.
- Created **Python scripts** to automate **data pipelines**, streamlining data from multiple sources to dashboards.
- **Managed the "Contact HP Support"** project on the self-service digital support website, including stakeholder meetings and coordination.

### HP Inc

*Data Analyst Intern*

Bangalore, India

Feb 2021 – Aug 2021

- Developed an Employee Experience **Game proof of concept** using React and Node.
- **Analyzed digital self-serve** customer experience data.

## Projects

### CX Copilot (AI Analyst)

Designed a full stack chatbot application to handle inquiries about CX data by **performing in-depth analysis** and mapping relevant nodes. The project involved automating complex data pipelines with **ETLs, fine-tuning and training open-source LLM models** in Databricks, and hosting the application in a Kubernetes cluster using Docker containerization.

### Root cause finder & Alert system Automation

An automation system that can perform data analysis, anomaly detection, and **root cause identification**. The script performs statistical analysis, **generates PowerPoint presentations** for visual summaries, and sends the reports to stakeholders.

### Python Automation for Data Pipeline Streaming (with Auto Adapting Capabilities)

Developed a Python automation script for performing Extract, Transform, Load (ETL) operations from various sources, streamlining data to dashboards. Designed with **auto-adapting capabilities** to accommodate changes in data structure seamlessly.

### HMTC (Hierarchical Multilabel Text Classification)

Implemented an HMTC model that provides deep insights in a **visually comprehensible 3-dimensional tree shape**, aiding stakeholders in effectively identifying root causes

## Skills

*Python, Data Science, Machine Learning, NLP/ LLM, Azure, Apache spark Databricks, ML Ops, Dev Ops, Statistics, Automations, Agile, Problem Solving, Docker.*

## Education

**College:** Pondicherry Engineering College: **Bachelor of Technology in Information Technology**

**CGPA:** 9.3/ 10.0

**HSC:** 95%

**SSLC:** 98%